A Genome-wide Association Study of Circulating Levels of Atorvastatin and Its Major Metabolites

Richard M. Turner[1,*], Vanessa Fontana[1], Jieying E. Zhang[1], Daniel Carr[1], Peng Yin[2,3], Richard FitzGerald[1], Andrew P. Morris[2,4] and Munir Pirmohamed[1]

Atorvastatin (ATV) is frequently prescribed and generally well tolerated, but can lead to myotoxicity, especially at higher doses. A genome-wide association study of circulating levels of ATV, 2-hydroxy (2-OH) ATV, ATV lactone (ATV L), and 2-OH ATV L was performed in 590 patients who had been hospitalized with a non-ST elevation acute coronary syndrome 1 month earlier and were on high-dose ATV (80 mg or 40 mg daily). The *UGT1A* locus (lead single nucleotide polymorphism, rs887829) was strongly associated with both increased 2-OH ATV/ATV ($P = 7.25 \times 10^{-16}$) and 2-OH ATV L/ATV L ($P = 3.95 \times 10^{-15}$) metabolic ratios. Moreover, rs45446698, which tags *CYP3A7*1C*, was nominally associated with increased 2-OH ATV/ATV ($P = 6.18 \times 10^{-7}$), and *SLCO1B1* rs4149056 with increased ATV ($P = 2.21 \times 10^{-6}$) and 2-OH ATV ($P = 1.09 \times 10^{-6}$) levels. In a subset of these patients whose levels of ATV and metabolites had also been measured at 12 months after hospitalization ($n = 149$), all of these associations remained, except for 2-OH ATV and rs4149056 ($P = 0.057$). Clinically, rs4149056 was associated with increased muscular symptoms (odds ratio (OR) 3.97; 95% confidence interval (CI) 1.29–12.27; $P = 0.016$) and ATV intolerance (OR 1.55; 95% CI 1.09–2.19; $P = 0.014$) in patients ($n = 870$) primarily discharged on high-dose ATV. In summary, both novel and recognized genetic associations have been identified with circulating levels of ATV and its major metabolites. Further study is warranted to determine the clinical utility of genotyping rs4149056 in patients on high-dose ATV.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ There is high interindividual variability in circulating atorvastatin (ATV) levels, and higher statin exposure increases the risk of statin-related muscle toxicity.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ What genomic variants are associated with circulating levels of ATV and its metabolites in patients using a genome-wide association study, and what is the clinical impact of identified variants?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ This study found the *UGT1A* locus to be strongly associated with ATV disposition. The role of *SLCO1B1* rs4149056 (c.521T>C, p.V174A) on increasing ATV exposure was confirmed, and *CYP3A7* was newly associated with ATV hydroxylation. Patients were primarily on high-dose ATV, and *SLCO1B1* rs4149065 increased the risk of muscular complaints and ATV intolerance.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

☑ The clinical impact of *SLCO1B1* rs4149056 in patients on high-dose ATV needs further study and, if confirmed, consideration for clinical implementation. Further research is required to understand the complex effects of *UGT1A* on ATV disposition and to study the influence of *CYP3A7* on other CYP3A-substrate drugs.

Statins are the cornerstone of cholesterol-lowering pharmacotherapy and among the most frequently prescribed medications worldwide. Atorvastatin (ATV) is the guideline-recommended first-line lipid-lowering drug in primary and secondary prevention of cardiovascular disease.[1]

The metabolism of ATV is shown in **Figure 1**. Briefly, ATV can be hydroxylated by cytochrome P450 3A (CYP3A) to 2-hydroxy (2-OH) ATV and 4-OH ATV, or undergo lactonization via an unstable acyl glucuronide intermediate to ATV lactone (ATV L).[2] ATV L can be similarly hydroxylated by CYP3A,[3] or 2-OH ATV and 4-OH ATV conceivably undergo lactonization, producing 2-OH ATV L and 4-OH ATV L. The lactone metabolites can be hydrolyzed via plasma paraoxonases to their corresponding hydroxy acids.[4]

[1]Department of Molecular and Clinical Pharmacology, University of Liverpool, Liverpool, UK; [2]Department of Biostatistics, University of Liverpool, Liverpool, UK; [3]Chinese Academy of Sciences, Shenzhen Institutes of Advanced Technology, Shenzhen, China; [4]Division of Musculoskeletal and Dermatological Sciences, University of Manchester, Manchester, UK. *Correspondence: Richard M. Turner (Richard.Turner@liverpool.ac.uk)

Received November 13, 2019; accepted January 28, 2020. doi:10.1002/cpt.1820

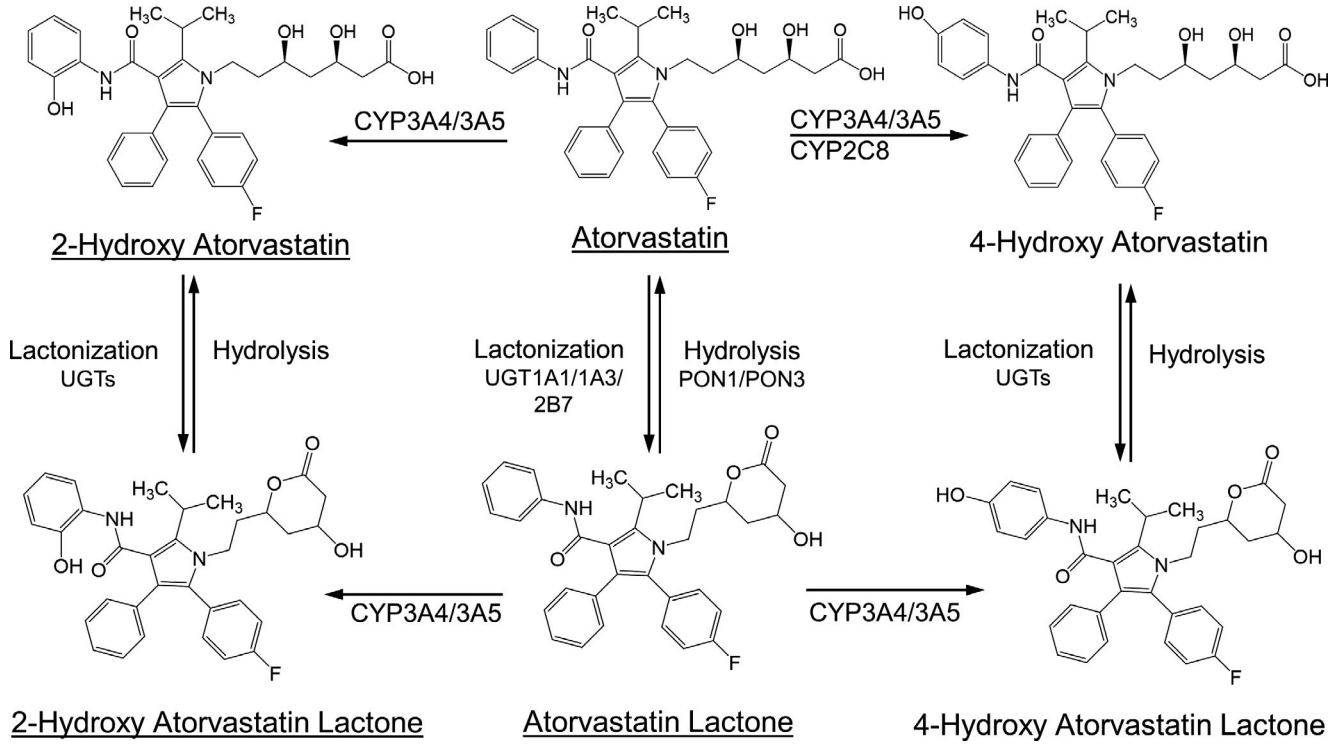

Figure 1 Atorvastatin (ATV) biotransformation. ATV can be hydroxylated by cytochrome P450 3A (CYP3A) to 2-hydroxy (2-OH) ATV and 4-OH ATV, or undergo lactonization via an unstable acyl glucuronide intermediate to ATV lactone (ATV L).[2] The main CYP3A enzyme responsible for ATV hydroxylation is CYP3A4 rather than CYP3A5, although CYP2C8 contributes a small amount to ATV 4-hydroxylation *in vitro*.[3] Uridine 5′-diphospho (UDP)-glucuronosyltransferase (UGT) 1A3 has the highest *in vitro* rate of ATV lactonization, followed by UGT1A1 and then UGT2B7.[39] Moreover, ATV L can be hydroxylated by CYP3A, or 2-OH ATV and 4-OH ATV conceivably undergo lactonization, producing 2-OH ATV L and 4-OH ATV L. ATV and its hydroxy-metabolites collectively inhibit 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase (HMGCR) to reduce circulating low-density lipoprotein cholesterol. The lactone metabolites are inactive against HMGCR, but can be hydrolyzed via plasma paraoxonases (PONs) to their corresponding hydroxy acids.[4] The major analytes are underlined and were investigated here.

Although generally well tolerated, statins are associated with adverse drug reactions in a small subset of patients, including statin-related myotoxicity (SRM) and new-onset diabetes mellitus.[5] SRM ranges from common muscular symptoms (~ 5% patients) where causal inference can be challenging to uncommon myopathies (~ 0.1%) and, rarely, rhabdomyolysis (0.1–8.4/100,000 patient-years).[5] Risk factors include higher statin dose, comedications that inhibit cytochrome P450 3A (CYP3A), and potentially increased circulating levels of statin lactone species, which are considered more myotoxic.[5-7] The solute carrier organic anion transporter family member 1B1 gene (*SLCO1B1*) encodes the hepatic xenobiotic influx transporter, OATP1B1. Importantly, a common nonsynonymous variant in *SLCO1B1*, rs4149056 (c.521T>C, p.V174A), is associated with elevated systemic exposure of ATV and all other statins except fluvastatin.[8] Furthermore, *SLCO1B1* rs4149056 has been consistently associated with simvastatin-related myotoxicity.[7,9]

Taking all the evidence together, increased systemic statin exposure seems to predispose to SRM. Therefore, identifying factors that alter ATV exposure is important, particularly given the high interindividual variability (45-fold) in circulating ATV levels that has been reported.[10] Candidate gene studies have reported variants affecting ATV pharmacokinetics in *ABCG2*,[11] *CYP3A4*,[12] *CYP3A5*,[12] and *PPARA*,[13] but to date no comprehensive genome-wide association study (GWAS) has been undertaken. Therefore, the aim of this study was to conduct a large GWAS of steady-state plasma levels of ATV and its major metabolites in a patient cohort on ATV following a recent non-ST elevation acute coronary syndrome (NSTE-ACS) and to relate the identified variants to clinical outcomes.

METHODS

Pharmacogenetics of acute coronary syndrome study

This investigation utilized the pharmacogenetics of acute coronary syndrome (PhACS) study, described previously.[14] Briefly, PhACS was a UK multicenter prospective observational study that recruited 1,470 patients hospitalized with an NSTE-ACS. Follow-up was at 1 (visit 2, V2) and 12 (V3) months, and annually thereafter until all patients had been followed up for at least 12 months. Patient demographics, comorbidities, medication information, and biosamples (blood and urine) were collected at recruitment; drug use, adherence, new events, and biosamples (V2/V3 only), were collected at follow-up. Participants were genotyped using the Illumina HumanOmniExpressExome-8 version 1.0 BeadChip at Edinburgh Genomics (Roslin Institute, Scotland).

The protocol was approved by the Liverpool (adult) research Ethics Committee (UK), site-specific approval was granted at all study sites, and informed consent was ascertained from all study subjects in accordance with the Declaration of Helsinki.

Determination of ATV and metabolite concentrations

The concentrations of ATV and its three major metabolites (2-OH ATV, ATV L, and 2-OH ATV L)[11] were quantified in EDTA plasma

samples from V2 (and V3 if eligible—see cohort selection below) using a high-performance liquid chromatography-tandem mass spectrometry assay developed and validated for this study.[15] The dynamic range for the analytes was 0.5–125 ng/mL. Briefly, each 50 μL aliquot of participant plasma was mixed with 20 μL composite internal standard solution containing a deuterated internal standard for each analyte, before 180 μL 100% acetonitrile containing 0.3% acetic acid was added. After centrifugation, the supernatant was diluted 1:2 in water and a 5 μL aliquot was injected into a 2.7 μm Halo C18 column (50 × 2.1 mm ID, 90Å; Hichrom Limited, Reading, UK) that separated analytes using gradient elution. Analyte detection was via multiple reaction monitoring (Sciex triple quadrupole 6500 QTRAP mass spectrometer with a Turbo V electrospray source; AB Sciex, Warrington, UK) using transitions (all $(M + H^+)$ *m/z*): ATV 559.3 → 440.3, 2-OH ATV 575.3 → 250.2, ATV L 541.3 → 276.2, 2-OH ATV L 557.3 → 276.2, ATV-d5 564.1 → 445.1, 2-OH ATV-d5 580.2 → 255.1, ATV-d5 L 546.3 → 281.2, and 2-OH ATV-d5 L 562.2 → 281.1.

GWAS cohort selection and end points

The inclusion criteria for the main cohort used for the GWAS of ATV and metabolite levels at V2 were: PhACS participants that had passed all genetic quality control (QC; see below) procedures, were taking ATV at the same dose of 80 or 40 mg daily at baseline and V2 (so steady-state had been reached), and both ATV adherence data and an EDTA sample were available at V2. Exclusion criteria were low outlier participants with two or more analytes < 0.5 ng/mL (i.e., less than the lower limit of quantification), and high outlier participants with ATV ≥ 31 ng/mL or ≥ 15.5 ng/mL if prescribed ATV 80 mg or 40 mg, respectively.

Participants with 2 or more analytes < 0.5 ng/mL were excluded because they were more likely to be statin nonadherent (reported missing at least one ATV pill in the previous week) compared with those with no low concentrations ($P = 0.001$). Participants with only 1 analyte < 0.5 ng/mL seemed similarly statin adherent compared with those with no low concentrations ($P = 0.88$), could have unusual ATV pharmacokinetics, and so were retained. Time since last dose was determined by assuming ATV was administered at 22:00 the preceding night, because only time of blood sampling was precisely known. An approximate time since last dose is a common limitation in observational pharmacogenomic studies,[16] and this time data cannot be retrospectively collected. The mean sampling time of 14 hours is also on the low curvature aspect of the analyte concentration-time profiles[17] and so the effect of sampling time measurement error is expected to be low.[16] Nevertheless, to minimize any high outlier effect from recent ATV administration, the ceiling of ≥ 31 ng/mL for ATV 80 mg was based *a priori* on mean ATV concentration plus 3 SDs from a previous smaller study that measured patient ATV levels.[10] This ceiling was halved for ATV 40 mg. Accordingly, time since last dose was reassuringly strongly associated with ATV levels here ($P = 4.20 \times 10^{-14}$).

All eligible participants were included in the main cohort to maximize discovery sample size. However, in a subset of these participants (cohort 2), ATV and metabolite concentrations were also determined at V3 to assess whether biologically plausible variants (see below) identified in the main cohort (at V2) remained associated at 12 months. These participants were selected if they were on ATV 80 mg daily at V2 and V3, had been adherent to ATV 80 mg at V2 and V3, and an EDTA sample was available at V3. ATV analyte quantification at V3 was limited to ~ 25% of those in the main cohort due to practical constraints. The same exclusion criteria were applied to the V3 concentrations as to the V2 concentrations in the main cohort.

The GWAS end points were: The individual concentrations of ATV and its three metabolites, analyte ratios (hydroxylation—2-OH ATV/ATV, 2-OH ATV L/ATV L; lactonization—ATV L/ATV), and the sum of analytes (ATV + 2-OH ATV + ATV L + 2-OH ATV L) at V2.

Cohort selection and end points for assessment of identified variants with clinical outcomes

PhACS participants that had passed genetic QC (see below) and were discharged from their baseline NSTE-ACS hospitalization on ATV at any dose formed a larger clinical cohort. This cohort was used to investigate the association between biologically plausible variants and time to major adverse cardiovascular (CV) events (MACEs; a composite of myocardial infarction, ischemic stroke, or CV death) and all-cause mortality (ACM). From this cohort, patients whose statin status (on statin or discontinuation), statin prescription, statin adherence, and adverse events were all known at V2 were used to assess whether identified variants were associated with any adverse event reported at V2 and attributed by the patient to their statin, and specifically muscular symptoms while on statin therapy, and ATV intolerance. ATV intolerance was defined as ATV discontinuation, ATV dose reduction, switching to a different statin of lower equivalent dose, and/or ATV nonadherence (self-reported missing at least one ATV pill in the last week) at V2.

Statistical analysis

Genetic quality control and imputation. There were 1,442 participants successfully genotyped. Per-individual QC was carried out excluding participants with genotype call rate < 95% ($n = 54$), discordant clinical/X-chromosome-derived sex ($n = 7$), and aberrant heterozygosity ($n = 1$). For each pair of individuals with an identity by descent > 0.1875 using pruned genetic data, the individual with the worse call rate and/or absence of CV events during follow-up was removed ($n = 5$). Potential confounding due to population stratification was assessed by principal component (PC) analysis (PCA). The reference PCA model was built using international HapMap 3 data from European (CEU), Asian (CHB + JPT), and African (YRI) ancestry individuals (**Figure S1**).[18] After application of the PCA model to PhACS, genetically non-European ancestry participants were excluded ($n = 18$), leaving $n = 1,357$ participants. Per-marker QC excluded single nucleotide polymorphisms (SNPs) with call rate < 95%, minor allele frequency < 5%, and SNPs deviating from Hardy-Weinberg equilibrium ($P < 0.0001$). QC was carried out in PLINK version 1.07 (http://zzz.bwh.harvard.edu/plink/).[19] Subsequently, the genotype scaffold was prephased using SHAPEIT version 2 (https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html)[20] and imputed to the 1000 Genomes Phase I reference panel (all ancestries, March 2012 release) using IMPUTE2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html).[21] Post-GWAS QC excluded SNPs with information score (a measure of imputation quality) < 0.4 or minor allele frequency < 1%.

GWAS. The ATV analytes, analyte ratios, and sum total were $\log_{10}$ transformed to adjust for right skew. Prior to GWAS, multivariable linear regression was used for covariate selection (**Tables S1-S3**), as done previously with sparse ATV sampling.[10] Clinical variables with a univariate $P \leq 0.1$ underwent multivariable linear regression modeling using stepwise selection to determine the baseline clinical covariate model for each end point ($P < 0.05$ taken to indicate clinical covariate multivariable significance). All clinical variables in the main cohort had < 2% data missing. The clinical variables selected were very similar to those described previously.[22] To adjust for fine-scale European ancestry population structure, the first two principal components were also included as covariates in all GWAS analyses.

For each end point, a GWAS was conducted using frequentist association testing assuming an additive model of SNP effect and considering genotype dosages within SNPtest version 2.5 (https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html).[23] Manhattan and QQ plots were created using the qqman package[24] within the R statistical framework[25]; lambda genomic control inflation factor was also calculated. A Bonferroni multiple testing-corrected genome-wide statistical significance threshold of $P < 5.0 \times 10^{-8}/8$ tests $= 6.25 \times 10^{-9}$ was applied.

Appraisal of genomic signals. Genomic signals were investigated using: Regional locus plots (http://locuszoom.sph.umich.edu/),[26] Ensembl Variant Effect Predictor (http://www.ensembl.org/info/docs/tools/vep/index.html),[27] expression quantitative trait loci (eQTL) using the Genotype-Tissue-Expression analysis release V8 (https://www.gtexportal.org/home/),[28] and assessment based on prior knowledge of gene action and statin pharmacology. From the above, functional variants with biological plausibility were identified. These variants were conditioned on, alongside the covariates, and their region re-analyzed within SNPtest version 2.5. Lastly, the proportion of observed variability (R^2) explained by these variants was assessed by multivariable linear regression.

Assessment of biologically plausible variants with analyte levels at 12 months. Identified variants were further assessed by determining their association to their respective V3 analyte end point(s) by multivariable linear regression within cohort 2, adjusted for age, sex, time since last V3 ATV dose, and sample storage duration ($P < 0.05$ taken to be significant) if included in an end point's V2 covariate model (**Table S3**).

Sensitivity analysis. A localized chromosome 2 analysis of 2-OH ATV/ATV was conducted after exclusion of all outlying participants whose 2-OH ATV/ATV ratio was outside two SDs from the mean to investigate whether the locus signal was attributable to outliers.

Clinical end point analysis. Cox proportional hazards regression was used in the clinical cohort to assess the impact of biologically plausible SNPs on time to MACE and ACM during all follow-up from baseline discharge. For MACE, participants were censored at the earliest of the date of non-CV disease death or date of last recorded visit. For ACM, participants were censored at the date of the last recorded visit. The clinical end points determined at V2 (statin adverse events, muscular symptoms, and ATV intolerance) were analyzed by logistic regression. In all clinical analyses, if a SNP had univariate $P \leq 0.1$, its adjusted association was tested by adding it to a multivariable model containing clinical covariates (with univariate $P \leq 0.1$) that had been chosen by forward likelihood ratio selection. SNPs were tested using an additive model, considering genotype dosage when necessary. In addition, the borderline association between *SLCO1B1* rs4149056 and muscular complaints was explored further using a dominant model given the low number of reported muscle complaints and its role in simvastatin myotoxicity.[7] The multiple testing-corrected significance threshold was $P < 0.05/16$ tests $= 0.003$.

The 12 month and clinical analyses were conducted using IBM SPSS version 22.0 (IBM, Armonk, NY).

RESULTS

GWAS

Overall, 590 patients were included in the main cohort. The study selection process is summarized in **Figure 2**. In the main cohort, 77% of patients were men, median age was 64 years old, 551 were on ATV 80 mg, and 39 on ATV 40 mg daily. The median concentrations of ATV 80 mg and 40 mg were 5.2 and 4.4 ng/mL, respectively. The clinical characteristics of the main cohort, and clinical variables selected for the analyte multivariable covariate models are in **Supplementary Tables S1** and **S3**, respectively. Following QC, the number of SNPs preimputation and postimputation were 598,054 and 8,659,258, respectively. The Manhattan plots for ATV, ATV L, 2-OH ATV, and 2-OH ATV L are presented in **Figure 3**, and the plots for 2-OH ATV/ATV, 2-OH ATV L/ATV L, ATV L/ATV, and the sum of analytes are in **Figure 4**. Importantly, the *UGT1A* locus (lead SNP, rs887829) on chromosome 2 was associated with increased metabolic ratios of both 2-OH ATV/ATV ($P = 7.25 \times 10^{-16}$) and 2-OH ATV L/ATV L ($P = 3.95 \times 10^{-15}$) at genome-wide significance (**Figure 4**, **Table 1**). The untransformed, unadjusted median 2-OH ATV/ATV ratio increased from 1.1 to 1.3 to 1.8 in those with none, 1 or 2 rs887829 minor alleles, respectively. The lead SNP, rs887829, was genotyped on the array and its minor allele was an eQTL within the Genotype-Tissue-Expression project associated with higher *UGT1A3* expression and lower *UGT1A1* expression in liver, and higher *UGT1A7* and lower *UGT1A6* expression in esophagus mucosa. No other genome-wide significant loci were identified. There was no evidence of genomic inflation (**Figure S2**).

Biologically plausible signals

From the GWAS analysis, three biologically plausible loci were identified: the genome-wide significant *UGT1A* locus, and two further loci, *SLCO1B1* and *CYP3A7*, of nominal significance (**Table 1**). **Figure 5** shows regional plots of these loci. The full list of nominally associated lead SNPs are available in **Table S4**.

The ATV chromosome 12 lead SNP (rs4363657; $P = 1.63 \times 10^{-6}$) was in *SLCO1B1* and was in linkage disequilibrium (LD; $r^2 = 0.70$) with the *SLCO1B1* missense SNP, rs4149056 (p.V174A), known to increase ATV levels[17]; both rs4363657 and rs4149056 were genotyped on the array. The rs4149056 minor allele was also nominally associated with higher ATV ($P = 2.21 \times 10^{-6}$) and 2-OH ATV ($P = 1.09 \times 10^{-6}$) concentrations, but not with ATV L ($P = 0.0032$) or 2-OH ATV L ($P = 0.0042$) levels, in keeping with previous healthy volunteer results.[17] The untransformed, unadjusted median ATV concentration increased from 4.7 to 6.3 to 10.9 ng/mL in those with none, 1 or 2 rs4149056 minor alleles, respectively.

The minor allele of a lead chromosome 7 SNP (rs45446698; T>G) was nominally associated with increased 2-OH ATV/AVT ($P = 6.18 \times 10^{-7}$) and 2-OH ATV L/ATV L ($P = 2.91 \times 10^{-5}$; **Table 1**) ratios. The untransformed, unadjusted median 2-OH ATV/ATV ratio increased from 1.2 to 1.8 in those with none or one rs45446698 minor allele, respectively; no individual was homozygous. SNP rs45446698 is an eQTL associated with higher CYP3A7 expression in the liver, terminal ileum of the small intestine, and adrenal gland.[29] Moreover, rs45446698 is one of seven correlated SNPs (rs11568824, rs45494802, rs45575938, rs45467892, rs11568825, rs11568826, and rs45446698) that cluster within the *CYP3A7* promoter and constitute the *CYP3A7*1C* allele, which has been associated with increased adult CYP3A7 expression.[30,31] Although rs45446698 was imputed (info score 0.93), rs11568824 was genotyped on the array, is in LD with rs45446698 ($r^2 = 0.74$), and was also correlated with increased 2-OH ATV/ATV ($P = 1.23 \times 10^{-5}$) and 2-OH ATV L/ATV L ($P = 6.31 \times 10^{-5}$) hydroxylation ratios.

In the sensitivity analysis that excluded patients with outlying 2-OH ATV/ATV ratios, there was minimal impact on the rs887829 association signal ($P = 7.25 \times 10^{-16}$ to $P = 3.31 \times 10^{-14}$), indicating that the *UGT1A* signal is not attributable to outliers. Conditioning separately on rs4149056 (*SLCO1B1*), rs45446698 (*CYP3A7*), and rs887829 (*UGT1A*), led to a complete loss of

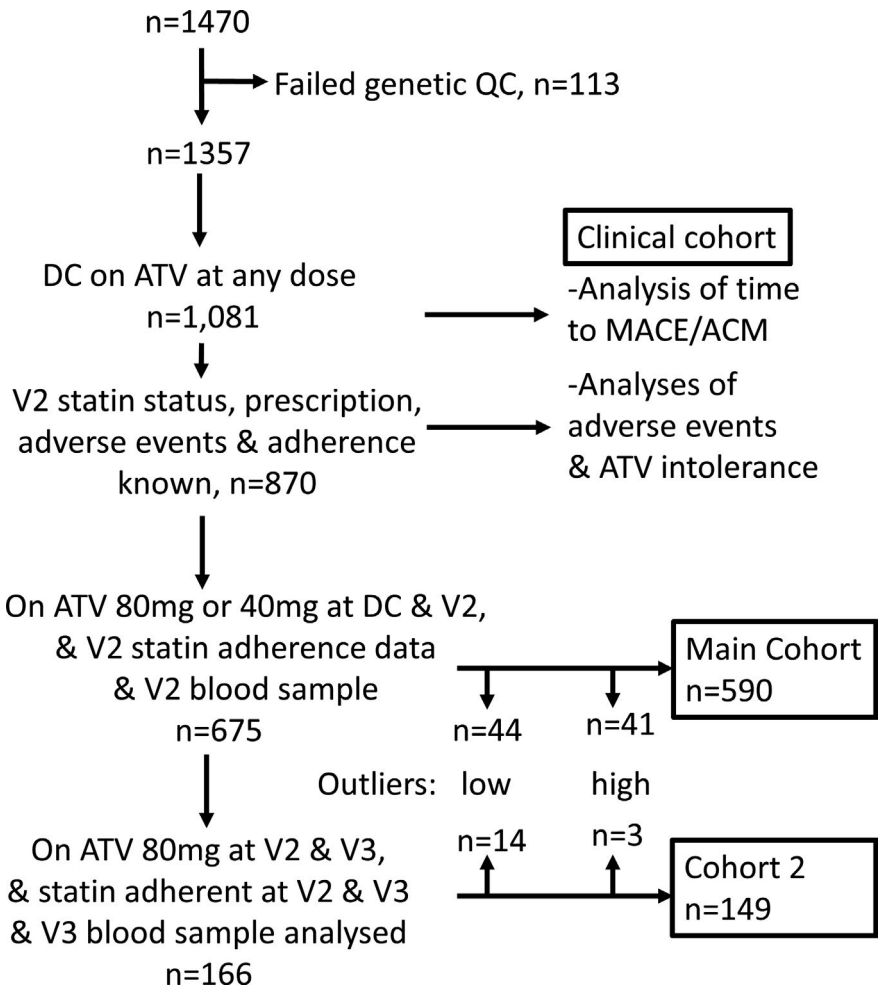

Figure 2 The study cohort selection process. Time points: DC = discharge from baseline hospitalization for a non-ST elevation acute coronary syndrome; V2 = visit 2 (1-month follow-up); V3 = visit 3 (12-month follow-up). The clinical end points of adverse events, muscular symptoms, and ATV intolerance required statin status (on statin, or discontinuation), prescription (type of statin, dose), patient-reported statin adverse events, and statin adherence to be all known at V2. Low outliers were participants with two analytes < 0.5 ng/mL (lower limit of quantification). High outliers were those with ATV ≥ 31 ng/mL (on ATV 80 mg) or ≥ 15.5 ng/mL (ATV 40 mg). MACE/ACM, major adverse cardiovascular event/all-cause mortality; QC, quality control.

signal for each end point. Inclusion of rs4149056 into the ATV model increased the proportion of ATV concentration variance explained from 16.9% to 20.2%; sequential addition of rs887829 and rs45446698 to the 2-OH ATV/ATV multivariable clinical model increased variability explained from 15.8% to 25.2% to 28.6%, respectively (**Table S5**).

Assessment of signals with analyte levels at 12 months

At V3, analyte concentrations from 149 patients were available (cohort 2). The associations remained significant for rs887829 (*UGT1A*) and rs45446698 (*CYP3A7*; **Table 1**). *SLCO1B1* rs4149056 was associated with ATV ($P = 0.0082$) but not 2-OH ATV concentration ($P = 0.057$; **Table 1**).

Clinical analyses

The discharge ATV doses and clinical end point constituents of the clinical cohort ($n = 1,081$) are summarized in **Table S6**. Briefly, 90% of patients were discharged on ATV 80 mg, 7% on ATV 40 mg, and the remainder on ATV 10–30 mg. There were 142 MACE and 93 ACM events. From this cohort, 870 had statin status, statin prescription, statin adherence, and adverse events all available at V2; of these: 53 reported any adverse event, 13 specifically muscular complaints, 118 were ATV intolerant, and 752 ATV tolerant. *SLCO1B1* rs4149056 was nominally associated with both muscular symptoms ($P = 0.016$) and ATV intolerance ($P = 0.014$; **Table 2**). However, rs4149056 was not associated with MACE or ACM, and no association surpassed the multiple testing threshold. Neither rs45446698 (*CYP3A7*) nor rs887829 (*UGT1A*) were associated with any clinical end point.

DISCUSSION

The main findings of this study were: Confirmation of the association of rs4149056 (*SLCO1B1*) with elevated ATV and 2-OH ATV systemic levels, identification of rs45446698 (*CYP3A7*) with ATV hydroxylation, and a single strong genome-wide significant signal between the *UGT1A* locus (rs887829) and both

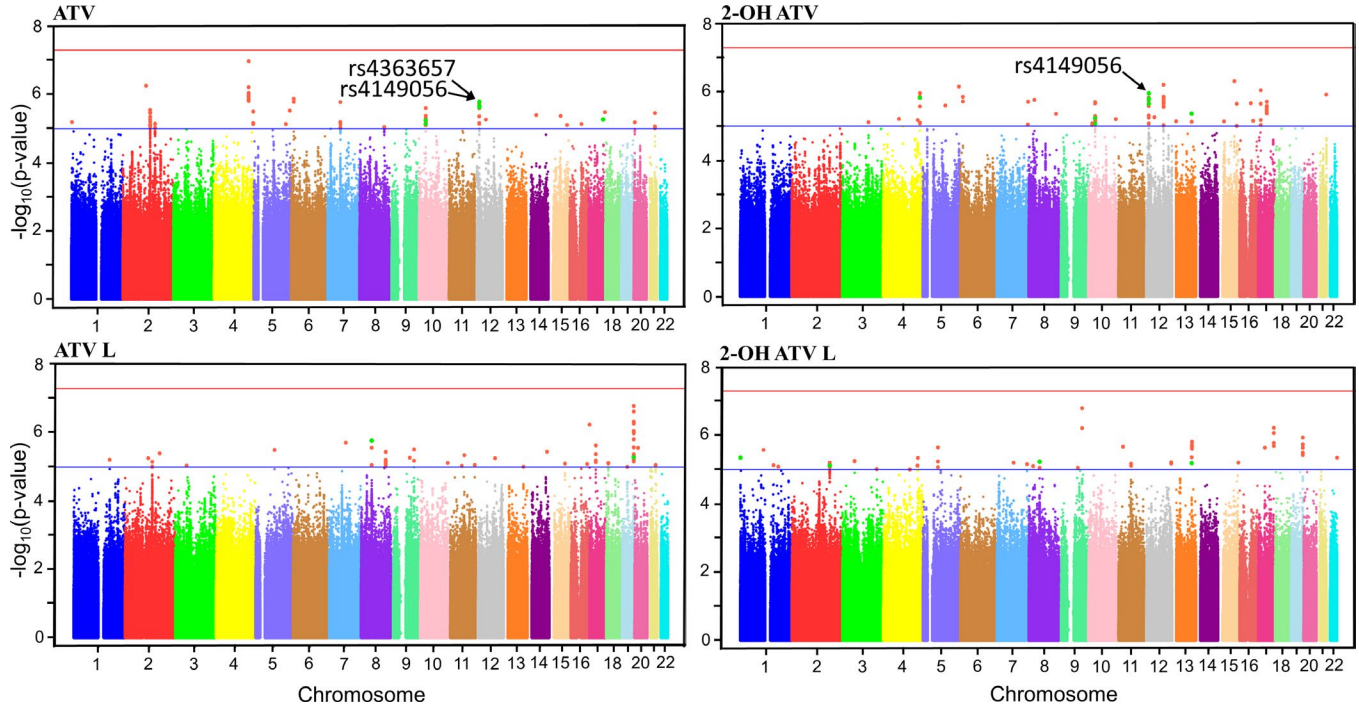

Figure 3 Manhattan plots of atorvastatin (ATV) and metabolites. Genome-wide association analyses were carried out using $\log_{10}$ transformed analyte concentrations in the main cohort, adjusted for the first two principal components and selected clinical covariates (as specified in **Table S3**), by frequentist association testing assuming an additive model of effect and considering genotype dosage. Green dots indicate genotyped variants.

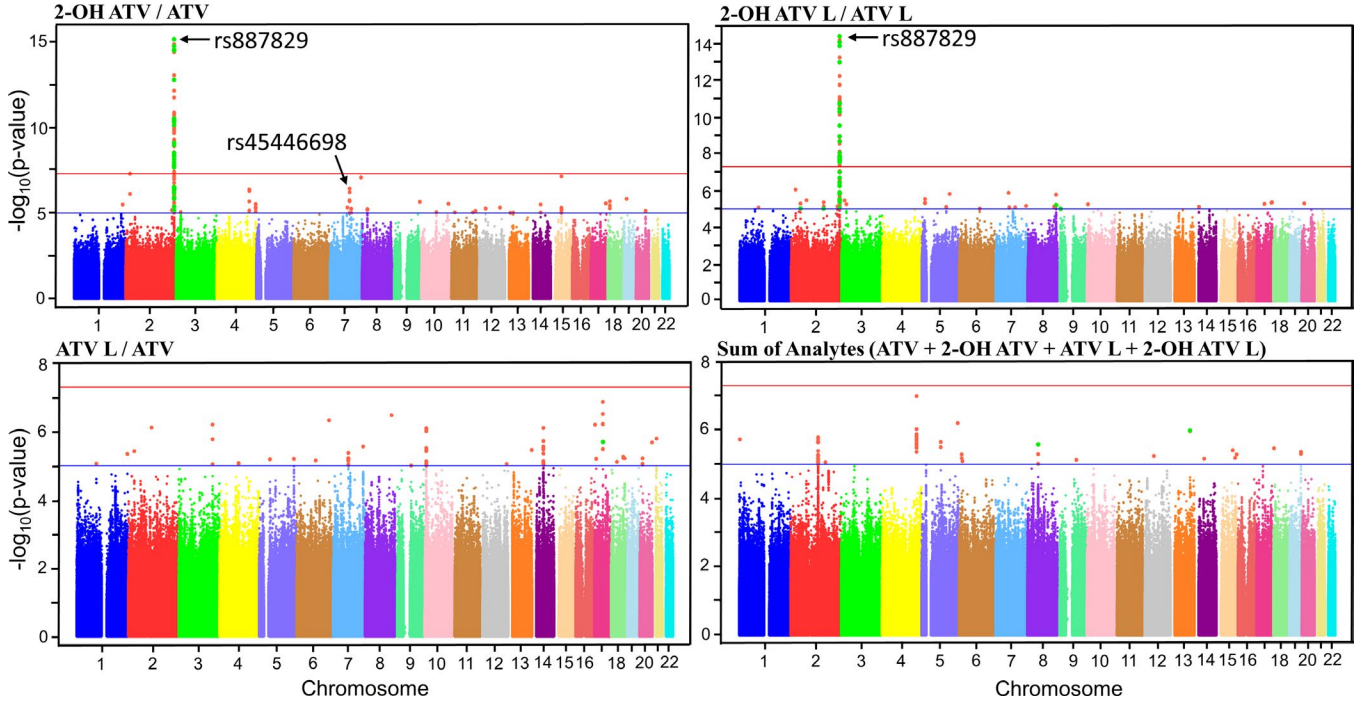

Figure 4 Manhattan plots of atorvastatin (ATV) metabolic ratios and the sum of analytes. Genome-wide association analyses were carried out using $\log_{10}$ transformed metabolic ratios or analyte concentrations within the main cohort, adjusted for the first two principal components and selected clinical covariates (**Table S3**), by frequentist association testing assuming an additive model of effect and considering genotype dosage. Green dots indicate genotyped variants.

2-OH ATV/ATV and 2-OH ATV L/ATV L metabolic ratios. The minor allele of rs4149056 was nominally associated with an increased risk of muscular symptoms and ATV intolerance.

The minor allele of nonsynonymous *SLCO1B1* rs4149056 is associated with decreased intrinsic OATP1B1 transport activity, a 221% increase in area under the simvastatin acid concentration-time

Table 1 Biologically plausible identified loci

Chr	Locus	SNP[a]	Position	Reference allele	Minor allele	Study MAF	1000G MAF[b]	Cohort[c]	Genome-wide association result[d]			
									Analysis	*n*[e]	B (SE)[f]	*P*-value
2	*UGT1A*	rs887829	233759924	C	T	0.35	0.30	Main	2-OH ATV/ATV	580	0.100 (0.012)	7.25×10^{-16}
									2-OH ATV L/ATV L	590	0.096 (0.012)	3.95×10^{-15}
								2	2-OH ATV/ATV	144	0.102 (0.024)	5.00×10^{-5}
									2-OH ATV L/ATV L	147	0.110 (0.024)	8.00×10^{-6}
7	5' of *CYP3A7*	rs45446698	99735325	T	G	0.05	0.04	Main	2-OH ATV/ATV	580	0.154 (0.031)	6.18×10^{-7}
									2-OH ATV L/ATV L	590	0.128 (0.030)	2.91×10^{-5}
								2	2-OH ATV/ATV	144	0.162 (0.048)	0.00104
									2-OH ATV L/ATV L	147	0.120 (0.048)	0.0140
12	*SLCO1B1*	rs4149056	21178615	T	C	0.15	0.16	Main	ATV	580	0.127 (0.026)	2.21×10^{-6}
									2-OH ATV	572	0.110 (0.022)	1.09×10^{-6}
								2	ATV	146	0.165 (0.061)	0.00817
									2-OH ATV	146	0.103 (0.054)	0.0568

1000 G, 1000 Genomes Project Phase III; 2-OH, 2-hydroxy; ATV, atorvastatin; ATV L, atorvastatin lactone; Chr, chromosome; MAF, minor allele frequency; SNP, single nucleotide polymorphism.
[a]SNPs selected were loci lead SNPs (rs887829, rs45446698) or functionally recognized SNPs within an identified locus (rs4149056). [b]1000 Genomes Project Phase III European MAFs. [c]Genome-wide association study (GWAS) in the main cohort (using visit 2 data from month 1), or candidate assessment in cohort 2 (using visit 3 data from 12 months). [d]GWAS results for relevant analyte end points are presented only. [e]Analysis cohort size varied slightly between endpoints due to adjustment for different clinical variables within the multivariable covariate model for each end point (see **Table S3**). [f]Coefficient (standard error) of the minor allele relative to the reference allele.

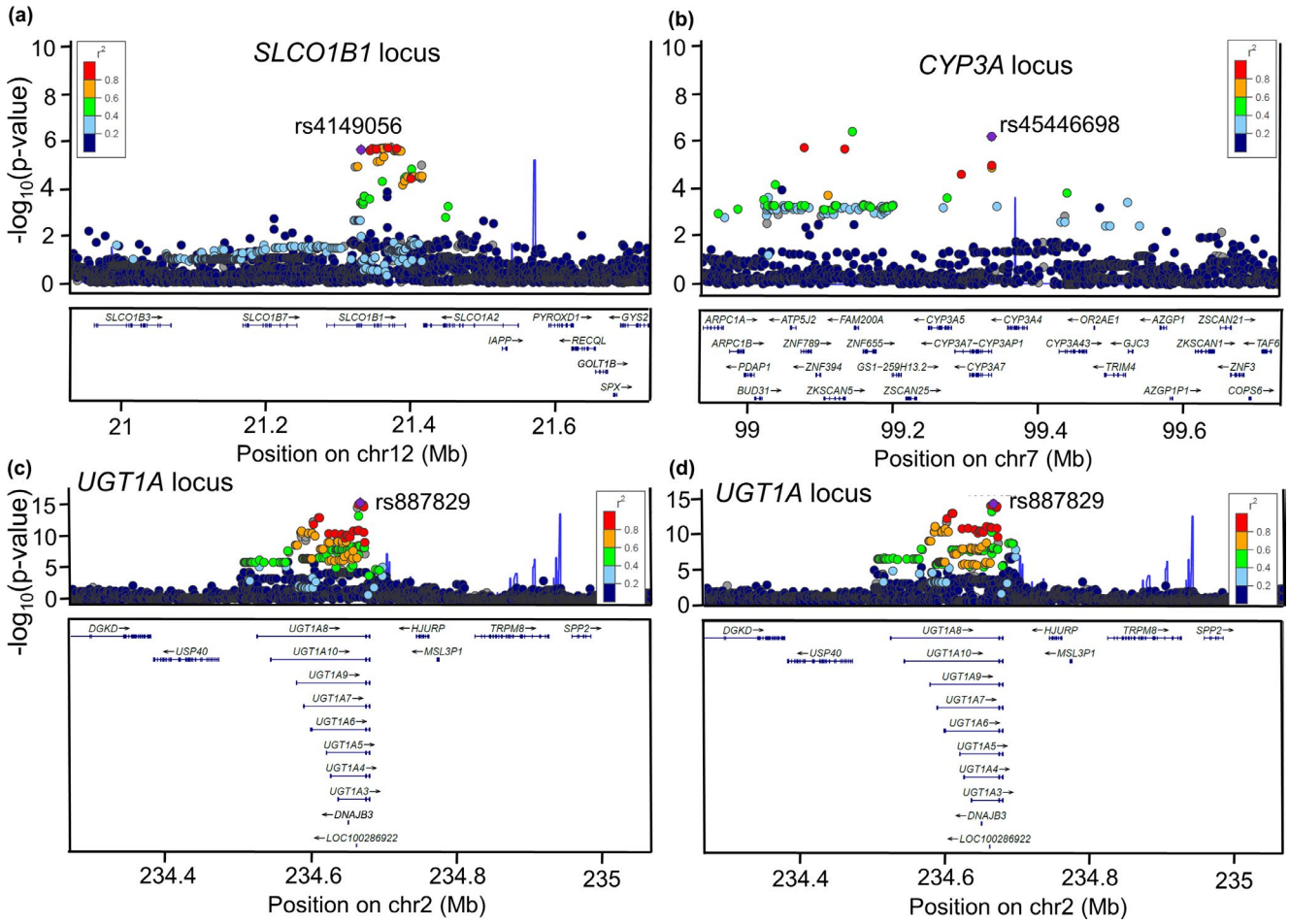

Figure 5 Regional plots of identified biologically plausible loci. Regional plots showing the *SLCO1B1* locus with atorvastatin (ATV) (**a**), the *CYP3A* locus, including *CYP3A7* with 2-OH ATV/ATV (**b**), and the *UGT1A* locus with 2-OH ATV/ATV (**c**), and 2-OH ATV L/ATV L (**d**). The color of the single nucleotide polymorphisms (SNPs) is based on linkage disequilibrium with the named SNP (in purple).

curve (AUC) in homozygotes, and simvastatin myotoxicity.[7,9,32] Carrying two rs4149056 minor alleles is also associated with 144% increase in ATV AUC in healthy volunteers.[17] However, rs4149056 was not significantly associated with ATV myotoxicity in a meta-analysis, although fewer participants took ATV relative to simvastatin in the included studies.[9] Interestingly, the rs4149056 minor allele has been associated with dose decreases, switching, and intolerance primarily to simvastatin (risk estimates 1.74–3.16[33,34]), and potentially in patients on ATV, but only if taking > 20 mg ATV daily (hazard ratio 3.26).[33] The current study is unusual in primarily investigating high-dose ATV (predominantly 80 mg), and adds further support that rs4149056 is associated with ATV intolerance. Moreover, statin discontinuation/nonpersistence/nonadherence are associated within an increased risk of MACE,[35] although no association between rs4149056 and MACE or ACM was observed here, perhaps due to the limited sample size.

The *UGT1A* locus at chromosome 2q37.1 contains 13 unique first exons, of which 4 are pseudogenes, followed by 4 common exons that can be spliced to the 9 functional first exons to give 9 alternate UGT1A phase II glucuronosyltransferases. The lead *UGT1A* SNP, rs887829, is noncoding and its variant allele has been associated in GWAS with elevated circulating bilirubin levels and cholelithiasis.[36] The variant allele of rs887829 is in tight LD with the dinucleotide tandem repeat reduction-of-expression *UGT1A1*28* allele in white patients,[37,38] which underpins Gilbert's syndrome (unconjugated hyperbilirubinemia), irinotecan toxicity, and increased risk of jaundice-associated atazanavir discontinuation.[38] Within the confines of this study, it was, however, not possible to genotype *UGT1A1*28*.

Statin acids are glucuronidated to an unstable acyl glucuronide intermediate that undergoes spontaneous cyclization to the corresponding statin lactone.[2] *In vitro*, UGT1A3 has the highest rate of ATV lactonization, followed by UGT1A1, whereas UGT2B7 catalyzes to a minor extent.[39] On the other hand, tissue expression of *UGT1A1* is higher than *UGT1A3* in liver and small intestine. Furthermore, *UGT1A3*2* is in strong LD with *UGT1A1**28. Thus, in one study, *UGT1A1*28* carriers had lower systemic ATV L exposure.[40] However, another study showed that *UGT1A3*2* is associated with increased hepatic microsomal UGT1A3 protein levels, increased *ex vivo* ATV lactonization, and increased ATV lactonization in healthy volunteers.[41] *UGT1A3*2*, *3, and *6 are haplotypes of rs2008584, rs1983023, and rs45449995, which were themselves associated with higher 2-OH ATV/ATV (P = 3.93 × 10^{-8}, 6.60 × 10^{-10} and 5.46 × 10^{-5}, respectively).

Table 2 Assessment of clinical impact of identified genetic variants

	All adverse events (n = 53)		Muscular symptoms (n = 13)		ATV intolerance (n = 118)[a]		MACE (n = 142 events)[b]		ACM (n = 93 events)[c]	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Univariate										
rs4149056	1.39 (0.68–2.26)	0.18	**2.32 (1.01–5.33)**	**0.048**[c]	**1.44 (1.02–2.03)**	**0.036**	1.08 (0.79–1.48)	0.64	1.29 (0.90–1.86)	0.17
rs4149056 (dominant)	–	–	**3.97 (1.29–12.27)**	**0.016**[c]	–	–	–	–	–	–
rs887829	0.90 (0.59–1.37)	0.61	0.45 (0.17–1.20)	0.11	0.98 (0.73–1.31)	0.89	1.04 (0.82–1.33)	0.74	1.25 (0.93–1.67)	0.14
rs45446698	0.62 (0.18–2.08)	0.43	0.92 (0.12–7.12)	0.94	0.76 (0.35–1.64)	0.49	0.66 (0.33–1.32)	0.25	0.40 (0.14–1.20)	0.10
Multivariable										
rs4149056	–	–	–	–	**1.55 (1.09–2.19)**[d]	**0.014**	–	–	–	–

The clinical cohort included all patients discharged from their index non-ST elevation acute coronary syndrome hospitalization on any ATV dose (n = 1,081) and was used to analyze time to MACE and time to ACM. The all adverse events, muscle symptoms, and ATV intolerance end points were analyzed in 870 of these patients in whom statin status (on statin, or discontinuation), prescription (statin, dose), adverse events, and statin adherence were all known at V2. Genotype analyses were additive unless otherwise specified. Bold text indicates nominally significant (P < 0.05) results.
ACM, all-cause mortality; ATV, atorvastatin; CI, confidence interval; HR, hazard ratio; MACE, major adverse cardiovascular event; OR, odds ratio.
[a]ATV intolerance composed of discontinuation (n = 23), ATV dose reduction (n = 26), statin switching to a lower equivalent dose (n = 22), and nonadherent to discharge ATV dose (n = 47). [b]Median follow-up from discharge was 17 months. [c]No clinical covariates were associated with muscular symptoms and so multivariable analysis not undertaken. [d]Adjusted for any $P2Y_{12}$ inhibitor and beta blocker at V2.

In exploratory analysis that adjusted for *UGT1A3*2*, **3*, and **6*, the signal was attenuated but not extinguished (rs887829 *P* value from 7.25×10^{-16} to 9.96×10^{-7} for 2-OH ATV/ATV), suggesting that *UGT1A* variants beyond common *UGT1A3* haplotypes and *UGT1A1*28* may also be involved.

Overall, the reason for the association here between rs887829 and increased hydroxylation, and not lactonization, is unclear. Interestingly, observed hydroxylation ratios for ATV and ATV L seem increased in *UGT1A3*2* carriers compared with wild-type healthy volunteers following ATV dosing.[42] Taking all the evidence together, reduced ATV lactonization with increased direct hydroxylation of ATV in rs887829 carriers seems perhaps less likely than increased ATV lactonization, with subsequent increased hydroxylation to 2-OH ATV L (and then higher 2-OH ATV levels following hydrolysis). The increased hydroxylation in the latter hypothesis is suggested by the rate of CYP-dependent metabolism of ATV L being 83-fold higher than for ATV.[3] However, further metabolic investigations will be required.

This, to the best of our knowledge, is the first patient study to implicate CYP3A7 in ATV metabolism by associating the rs45446698 minor allele with increased ATV hydroxylation. ATV is likely a CYP3A7 substrate[43] and has been shown *in vitro* to upregulate CYP3A7 protein expression.[44] Although the rs45446698 signal did not reach genome-wide significance ($P = 6.18 \times 10^{-7}$ for 2-OH ATV/ATV), it was also significant in the analysis of 12-month analyte levels ($P = 0.0010$). Neither of the common established reduction-of-function alleles, *CYP3A4*22* (rs35599367) and *CYP3A5*3* (rs776746), were in LD with rs45446698. Although *CYP3A4*22* was marginally associated with lower ratios of 2-OH ATV/ATV ($P = 0.016$) and 2-OH ATV L/ATV L ($P = 0.005$), it was not associated with ATV or ATV L concentrations; *CYP3A5*3* was not associated with the end points.

The human CYP3A subfamily consists of CYP3A4, 3A5, 3A7, and 3A43 located on chromosome 7q22.1.[45] CYP3A7 is the predominant CYP in fetal liver, accounting for 30–50% of total fetal liver CYP content.[46] However, the majority of total adult liver CYP3A content is CYP3A4.[45] Nevertheless, human liver *CYP3A7* mRNA expression varies > 700-fold,[47] and in ~ 10% of adult livers, CYP3A7 is present and contributes 9–36% of total CYP3A protein content.[48] In *CYP3A7*1C*, ~ 60 bp of the fetal *CYP3A7* promoter region is replaced by the corresponding region of the adult *CYP3A4* promoter.[49] Thus, *CYP3A7*1C* is associated with increased CYP3A7 mRNA expression in liver and intestine,[47] and likely higher liver protein levels.[48] Sequencing has previously confirmed that rs45446698 effectively tags *CYP3A7*1C*.[31] Clinically, rs45446698 is associated with progesterone and dehydroepiandrosterone sulphate levels.[50] Moreover, *CYP3A7*1C* has been associated with increased breast cancer mortality, all-cause mortality in lung cancer, and chronic lymphocytic leukemia progression, potentially due to increased metabolic deactivation of CYP3A-substrate chemotherapeutics.[31] Given that ~ 30% of clinically used drugs are metabolized by CYP3A,[45] the impact of *CYP3A7* variants warrant further investigation.

Study limitations

This study represents the first GWAS of ATV and metabolite levels. It was undertaken in a real-world patient cohort, and assessed the clinical impact of identified variants. However, external replication of the novel findings is required. Studies in patients of non-European ancestry are also required to determine generalizability. Genes with a recognized role in drug pharmacokinetics were focused on; thus, novel genes of nominal significance may have been detected (**Table S4**), but require functional characterization. Lastly, few patients self-reported muscular complaints and the etiology of muscular symptoms is difficult to establish; therefore, caution is required when interpreting the muscular symptoms analysis. Nevertheless, rs4149056 was also associated with ATV intolerance.

CONCLUSION

In summary, this GWAS has reaffirmed the impact of rs4149056 (*SLCO1B1*) on increasing ATV exposure and may also influence ATV intolerance and muscular complaints in patients on high-dose ATV. We have shown that *UGT1A* seems to be important in complex pathways associated with ATV disposition, whereas *CYP3A7* has newly been associated with increased ATV hydroxylation.

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

FUNDING

R.M.T. was supported by the North West England Medical Research Council (MRC) Training Scheme in Clinical Pharmacology and Therapeutics (G1000417). V.F. was supported by the Newton Fund and The Academy of Medical Sciences. The clinical study was funded as part of the National Health Service (NHS) Chair of Pharmacogenetics, Department of Health. M.P. is Emeritus NIHR Senior Investigator. The authors also thank the MRC Centre for Drug Safety Science for support.

CONFLICTS OF INTEREST

The authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS

All authors wrote the manuscript. R.M.T., M.P., R.F., and D.C. designed the research. R.M.T. and V.F. performed the research. R.M.T., P.Y., J.E.Z., and A.P.M. analyzed data.

1. NICE. Cardiovascular disease: risk assessment and reduction, including lipid modification (Clinical guideline 181) <https://www.nice.org.uk/guidance/cg181> (2016). Accessed October 23, 2019.
2. Prueksaritanont, T. *et al*. Glucuronidation of statins in animals and humans: a novel mechanism of statin lactonization. *Drug Metab. Dispos.* **30**, 505–512 (2002).
3. Jacobsen, W. *et al*. Lactonization is the critical first step in the disposition of the 3-hydroxy-3-methylglutaryl-CoA reductase inhibitor atorvastatin. *Drug Metab. Dispos.* **28**, 1369–1378 (2000).
4. Riedmaier, S., Klein, K., Winter, S., Hofmann, U., Schwab, M. & Zanger, U.M. Paraoxonase (PON1 and PON3) polymorphisms: impact on liver expression and atorvastatin-lactone hydrolysis. *Front. Pharmacol.* **2**, 41 (2011).
5. Alfirevic, A. *et al*. Phenotype standardization for statin-induced myotoxicity. *Clin. Pharmacol. Ther.* **96**, 470–476 (2014).
6. Skottheim, I.B., Gedde-Dahl, A., Hejazifar, S., Hoel, K. & Asberg, A. Statin induced myotoxicity: the lactone forms are more potent than the acid forms in human skeletal muscle cells in vitro. *Eur. J. Pharm. Sci.* **33**, 317–325 (2008).
7. Ramsey, L.B. *et al*. The clinical pharmacogenetics implementation consortium guideline for SLCO1B1 and simvastatin-induced myopathy: 2014 update. *Clin. Pharmacol. Ther.* **96**, 423–428 (2014).
8. Elsby, R., Hilgendorf, C. & Fenner, K. Understanding the critical disposition pathways of statins to assess drug-drug interaction risk during drug development: it's not just about OATP1B1. *Clin. Pharmacol. Ther.* **92**, 584–598 (2012).
9. Jiang, J. *et al*. Association between SLCO1B1 -521T>C and -388A>G polymorphisms and risk of statin-induced adverse drug reactions: a meta-analysis. *SpringerPlus* **5**, 1368 (2016).
10. DeGorter, M.K. *et al*. Clinical and pharmacogenetic predictors of circulating atorvastatin and rosuvastatin concentration in routine clinical care. *Circulat. Cardiovasc. Genet.* **6**, 400–408 (2013).
11. Keskitalo, J.E., Zolk, O., Fromm, M.F., Kurkinen, K.J., Neuvonen, P.J. & Niemi, M. ABCG2 polymorphism markedly affects the pharmacokinetics of atorvastatin and rosuvastatin. *Clin. Pharmacol. Ther.* **86**, 197–203 (2009).
12. Kitzmiller, J.P., Luzum, J.A., Baldassarre, D., Krauss, R.M. & Medina, M.W. CYP3A4*22 and CYP3A5*3 are associated with increased levels of plasma simvastatin concentrations in the cholesterol and pharmacogenetics study cohort. *Pharmacogenet. Genomics* **24**, 486–491 (2014).
13. Klein, K. *et al*. PPARA: a novel genetic determinant of CYP3A4 in vitro and in vivo. *Clin. Pharmacol. Ther.* **91**, 1044–1052 (2012).
14. Turner, R.M. *et al*. Investigating the prevalence, predictors, and prognosis of suboptimal statin use early after a non-ST elevation acute coronary syndrome. *J. Clin. Lipidol.* **11**, 204–214 (2017).
15. Turner, R.M., Fontana, V., Bayliss, M., Whalley, S., Santoyo Castelazo, A. & Pirmohamed, M. Development, validation and application of a novel HPLC-MS/MS method for the quantification of atorvastatin, bisoprolol and clopidogrel in a large cardiovascular patient cohort. *J. Pharm. Biomed. Anal.* **159**, 272–281 (2018).
16. Choi, L., Crainiceanu, C.M. & Caffo, B.S. Practical recommendations for population PK studies with sampling time errors. *Eur. J. Clin. Pharmacol.* **69**, 2055–2064 (2013).
17. Pasanen, M.K., Fredrikson, H., Neuvonen, P.J. & Niemi, M. Different effects of SLCO1B1 polymorphism on the pharmacokinetics of atorvastatin and rosuvastatin. *Clin. Pharmacol. Ther.* **82**, 726–733 (2007).
18. Altshuler, D.M. *et al*. Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).
19. Purcell, S. *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
20. Delaneau, O., Marchini, J. & Zagury, J.F. A linear complexity phasing method for thousands of genomes. *Nat. Meth.* **9**, 179–181 (2011).
21. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
22. Turner, R.M., Fontana, V., FitzGerald, R., Morris, A.P. & Pirmohamed, M. Investigating the clinical factors and co-medications associated with circulating levels of atorvastatin and its major metabolites in secondary prevention. *Br. J. Clin. Pharmacol.* **86**, 62–74 (2019).

23. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* **11**, 499–511 (2010).
24. Turner, S.D. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. bioRxiv <https://www.biorxiv.org/content/10.1101/005165v1> (2014). Accessed January 7, 2020.
25. R Core Team. R: a language and environment for statistical computing <https://www.R-project.org> (2016). Accessed January 7, 2020.
26. Pruim, R.J. *et al*. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).
27. McLaren, W. *et al*. The Ensembl variant effect predictor. *Genome Biol.* **17**, 122 (2016).
28. The GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
29. The GTEx Consortium. Multiple-tissue eQTL Comparison Plot for CYP3A7 99735325 <https://gtexportal.org/home/snp/rs45446698#eqtl-block> (2019). Accessed September 13, 2019.
30. Kuehl, P. *et al*. Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. *Nat. Genet.* **27**, 383–391 (2001).
31. Johnson, N. *et al*. Cytochrome P450 allele CYP3A7*1C associates with adverse outcomes in chronic lymphocytic leukemia, breast and lung cancer. *Cancer Res.* **76**, 1485–1493 (2016).
32. Pasanen, M.K., Neuvonen, M., Neuvonen, P.J. & Niemi, M. SLCO1B1 polymorphism markedly affects the pharmacokinetics of simvastatin acid. *Pharmacogenet. Genomics* **16**, 873–879 (2006).
33. de Keyser, C.E. *et al*. The SLCO1B1 c.521T>C polymorphism is associated with dose decrease or switching during statin therapy in the Rotterdam Study. *Pharmacogenet. Genomics* **24**, 43–51 (2014).
34. Donnelly, L.A. *et al*. Common nonsynonymous substitutions in SLCO1B1 predispose to statin intolerance in routinely treated individuals with type 2 diabetes: a go-DARTS study. *Clin. Pharmacol. Ther.* **89**, 210–216 (2011).
35. De Vera, M.A., Bhole, V., Burns, L.C. & Lacaille, D. Impact of statin adherence on cardiovascular disease and mortality outcomes: a systematic review. *Br. J. Clin. Pharmacol.* **78**, 684–698 (2014).
36. Milton, J.N. *et al*. A genome-wide association study of total bilirubin and cholelithiasis risk in sickle cell anemia. *PLoS One* **7**, e34741 (2012).
37. Horsfall, L.J. *et al*. Prevalence of clinically relevant UGT1A alleles and haplotypes in African populations. *Ann. Hum. Genet.* **75**, 236–246 (2011).
38. Gammal, R.S. *et al*. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for UGT1A1 and Atazanavir prescribing. *Clin. Pharmacol. Ther.* **99**, 363–369 (2016).
39. Schirris, T.J., Ritschel, T., Bilos, A., Smeitink, J.A. & Russel, F.G. Statin lactonization by uridine 5'-diphospho-glucuronosyltransferases (UGTs). *Mol. Pharm.* **12**, 4048–4055 (2015).
40. Stormo, C. *et al*. UGT1A1*28 is associated with decreased systemic exposure of atorvastatin lactone. *Mol. Diagn. Ther.* **17**, 233–237 (2013).
41. Riedmaier, S. *et al*. UDP-glucuronosyltransferase (UGT) polymorphisms affect atorvastatin lactonization in vitro and in vivo. *Clin. Pharmacol. Ther.* **87**, 65–73 (2010).
42. Cho, S.K., Oh, E.S., Park, K., Park, M.S. & Chung, J.Y. The UGT1A3*2 polymorphism affects atorvastatin lactonization and lipid-lowering effect in healthy volunteers. *Pharmacogenet. Genomics* **22**, 598–605 (2012).
43. DrugBank. Cytochrome P-450 CYP3A7 substrates <https://www.drugbank.ca/categories/DBCAT003800> (2019). Accessed September 17, 2019.
44. Weiß, F. *et al*. Indirect protein quantification of drug-transforming enzymes using peptide group-specific immunoaffinity enrichment and mass spectrometry. *Sci. Rep.* **5**, 8759 (2015).
45. Zanger, U.M. & Schwab, M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol. Ther.* **138**, 103–141 (2013).
46. Shimada, T. *et al*. Characterization of microsomal cytochrome P450 enzymes involved in the oxidation of xenobiotic chemicals in human fetal liver and adult lungs. *Drug Metab. Dispos.* **24**, 515–522 (1996).
47. Burk, O. *et al*. Molecular mechanisms of polymorphic CYP3A7 expression in adult human liver and intestine. *J. Biol. Chem.* **277**, 24280–24288 (2002).
48. Sim, S.C., Edwards, R.J., Boobis, A.R. & Ingelman-Sundberg, M. CYP3A7 protein expression is high in a fraction of adult human livers and partially associated with the CYP3A7*1C allele. *Pharmacogenet. Genom.* **15**, 625–631 (2005).
49. Gonzalez, F.J. The molecular biology of cytochrome P450s. *Pharmacol. Rev.* **40**, 243–288 (1988).
50. Ruth, K.S. *et al*. Genome-wide association study with 1000 genomes imputation identifies signals for nine sex hormone-related phenotypes. *Eur. J. Hum. Genet.* **24**, 284–290 (2016).